

# Evaluating AI-Powered Automation of Therapy Session Notes: A Pilot Randomized Controlled Trial

Anastasia Batkhina 

Europe and Central Asia Regional Office, UNICEF, Istanbul, Türkiye  
Email: batkhina.anastasia@gmail.com

**How to cite this paper:** Batkhina, A. (2025). Evaluating AI-Powered Automation of Therapy Session Notes: A Pilot Randomized Controlled Trial. *Psychology*, 16, 789-800.  
<https://doi.org/10.4236/psych.2025.166044>

**Received:** May 27, 2025  
**Accepted:** June 24, 2025  
**Published:** June 27, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Artificial intelligence (AI) has been increasingly integrated into various domains of healthcare, yet its application in psychotherapy remains underexplored. One of the most time-consuming tasks for psychotherapists is documentation, which can consume up to 20% of their working hours, contributing to burnout and reducing time spent on direct patient care. AI-powered tools designed to automate session note-taking offer a potential solution to this challenge. However, empirical evidence on their effectiveness is limited. The present study investigates the impact of Yung Sidekick, an AI-based documentation tool, on psychotherapists' administrative workload, adherence to treatment plans, and perceived therapy progress. A randomized controlled trial was conducted with 70 licensed psychotherapists in the United States. Participants were randomly assigned to an experimental group using Yung Sidekick for one month or a control group maintaining standard documentation practices. Outcome measures included time spent on session notes and preparation, adherence to treatment plans, therapy progress, and professional well-being indicators. Results demonstrated that the experimental group showed significant reductions in time spent on session notes and preparation, as well as improvements in adherence to treatment plans and perceived therapy progress. However, no significant changes were observed in other well-being measures such as professional stress and burnout. These findings suggest that AI-assisted documentation can enhance efficiency and adherence to structured treatment approaches, but further research is needed to examine long-term effects and potential placebo influences. Implications for integrating AI into psychotherapy practice and future research directions are discussed.

## Keywords

Psychotherapy, AI in Psychotherapy, Session Notes, Therapist Efficiency,

## 1. Introduction

A significant portion of a psychotherapist's workload involves taking detailed notes and writing reports, which is a time-consuming and often burdensome task (Budd, 2023). Psychotherapists in many countries face pressure to work within imperfect electronic health record (EHR) systems, which further complicate documentation processes (Dymek et al., 2021). Research has shown that therapists can spend up to 20% of their working hours on documentation (Budd, 2023). This administrative burden reduces the time available for direct patient care, contributes to therapist burnout, and lowers overall job satisfaction. Efficient and accurate documentation is crucial for maintaining high standards of care, tracking patient progress, and ensuring adherence to treatment plans (Ebberts et al., 2022). The challenge lies in streamlining this process without compromising the quality and depth of session records. Addressing this gap requires not only technology that integrates seamlessly into therapists' workflows but also solutions that demonstrate measurable improvements in efficiency, clinical outcomes, and overall therapist well-being.

AI-driven solutions have the potential to address these challenges by automating the documentation process for therapists (D'Alfonso, 2020; Ghassemi et al., 2020). These tools capture key aspects of therapy sessions, whether conducted online or in person, and extract relevant information such as topics, themes, symptoms, medications, and goals. By automating these tasks, AI aims to reduce the time therapists spend on note-taking while improving the accuracy and comprehensiveness of their documentation. Automation of routine administrative tasks has been shown to enhance therapist efficiency, reduce burnout, and improve job satisfaction (Ebberts et al., 2022), which may also lead to improved client outcomes and stronger long-term therapeutic relationships.

In related medical fields, AI tools have already proven effective. A recent meta-analysis demonstrated significant improvements in clinical documentation accuracy and efficiency through AI technologies, leading to a reduction in clinician workload and streamlined documentation processes (Lee, Britto, & Diwan, 2024). However, studies also indicate resistance to AI tools and slow adaptation in mental health care settings, often due to concerns over technological reliability, security, and the time required for training (Jacob, Sanchez-Vazquez, & Ivory, 2020; Zhang et al., 2023). These barriers highlight the need for further research into AI implementation strategies that support user trust and seamless integration into clinical practice.

While previous studies have explored the potential benefits of AI in streamlining medical documentation, limited research has investigated its direct impact on psychotherapy practice. The few existing studies suggest that AI solutions may

improve clinical efficiency by reducing the time clinicians spend on mental health assessments (Rollwage et al., 2023). AI has also been shown to enhance patient flow by optimizing administrative tasks and resource allocation in mental health practices (Dawoodbhoy et al., 2021). However, research has largely overlooked other potential applications of AI beyond these areas. This study aims to fill this gap by evaluating the effectiveness of the AI tool Yung Sidekick in automating therapy session documentation and examining its impact on key therapist-related outcomes, including time spent on administrative tasks, adherence to treatment plans, and perceived therapy progress. By providing empirical evidence on the role of AI in psychotherapy, this research contributes to the growing field of digital mental health innovation and its practical implications for clinical settings.

## 2. Present Research

The integration of AI tools like Yung Sidekick into psychotherapeutic practice presents a compelling opportunity to enhance therapist efficiency and job satisfaction. However, the adoption of such tools also raises important questions regarding their reliability, user acceptance, and overall impact on therapy outcomes. While AI has the potential to significantly alleviate the administrative burden on therapists, there is a need for empirical research to validate these benefits and address concerns regarding accuracy, usability, and therapist engagement with the technology. The current study builds upon prior findings by empirically evaluating the impact of Yung Sidekick in a randomized controlled trial. The existing literature highlights the critical role of technological interventions in reducing administrative burdens in healthcare (Philippe et al., 2022). By addressing these challenges, this research aims to contribute to the growing body of evidence supporting AI-driven innovations in mental health practice.

The primary objective of this study is to evaluate the effectiveness of Yung Sidekick in automating therapy session notes and its impact on various aspects of therapists' professional practice. Since there are currently few studies on such tools, this study is exploratory in nature, seeking to determine whether an AI tool can influence key aspects of psychotherapists' professional performance.

## 3. Method

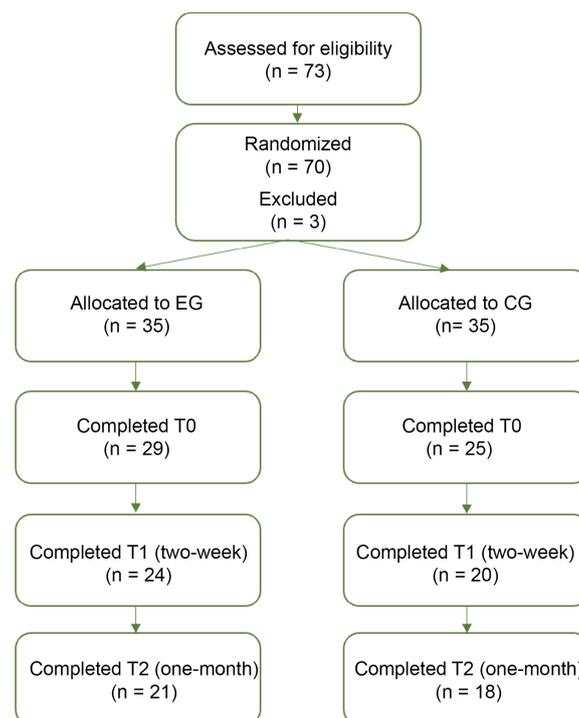
The study was preregistered on the Open Science Framework on July 12, 2024. This study is an experimental, pilot randomized control trial. It is designed to assess the effectiveness of the AI tool, Yung Sidekick, in automating therapy session notes and its impact on various aspects of psychotherapists' professional practice over time. The study data are available upon request from the corresponding author. No conflicts of interest are declared for this study.

### 3.1. Participants Characteristics

Participants were licensed psychotherapists practicing in the United States. Eligibility criteria included: (1) holding a valid psychotherapy license in the U.S., (2)

having at least one year of professional experience, (3) conducting at least 10 hours of therapy per week, and (4) no prior experience with AI-based note-taking tools. All participants provided informed consent prior to enrollment. The AI tool used in the experimental group was fully HIPAA-compliant.

As shown in **Figure 1**, a total of 73 participants completed the preliminary questionnaire, of whom 3 were excluded due to not meeting eligibility criteria. The remaining 70 participants were randomly assigned to the experimental ( $n = 35$ ) and control ( $n = 35$ ) groups. At baseline (T0), 29 participants in the experimental group and 25 in the control group completed assessments. At the two-week follow-up (T1), the sample included 24 and 20 participants, respectively. By the one-month follow-up (T2), 21 participants remained in the experimental group and 18 in the control group. The experimental group was 58% female, with an average age of 41.48 years ( $SD = 8.16$ ), 8.46 years ( $SD = 4.53$ ) of experience, and an average caseload of 33.07 clients ( $SD = 17.96$ ). The control group was 55% female, with an average age of 43.04 years ( $SD = 10.52$ ), 9.12 years ( $SD = 5.03$ ) of experience, and an average caseload of 31.15 clients ( $SD = 19.24$ ). Participants received a \$50 gift card upon study completion.



**Figure 1.** CONSORT flow diagram.

### 3.2. Sampling Procedure

Participants were recruited online through LinkedIn advertisements and direct outreach to therapy clinics and group practices. Randomization was conducted using a computer-generated sequence. No stratification was applied to ensure equal distribution of participant characteristics across conditions. Participants will be

aware of their group assignment (experimental or control) as it involves the use of the AI tool, which cannot be concealed. Researchers involved in the recruitment and initial interaction with participants will be aware of group assignments.

Based on the principles outlined by Whitehead et al. (2016) for determining sample sizes in pilot studies, a main trial designed to achieve 90% power with a two-sided 5% significance level would require sample sizes of 75, 25, 15, and 10 for treatment arms corresponding to extra small ( $\leq 0.1$ ), small (0.2), medium (0.5), and large (0.8) standardized effect sizes, respectively. Hence, the sample size used in this pilot study is deemed appropriate.

### 3.3. Measures

#### *Primary Outcomes*

1) *Time Spent on Session Notes*. Self-reported average time (in minutes) spent documenting notes for each therapy session.

2) *Time Spent on Session Preparation*. Self-reported average time (in minutes) spent preparing for each therapy session.

#### *Secondary Outcomes*

3) *Adherence to Treatment Plans*. Assessed using a standardized adherence checklist: “How often do you follow the prescribed treatment plans for your clients?”, 1-never, 5-always.

4) *Therapist Self-Efficacy*. Measured with the Therapist Self-Efficacy Scale (Gori et al., 2022;  $\alpha = 0.87$ ), a 21-item measure using a 5-point Likert scale, e.g. “During psychological treatment or psychotherapy sessions, I am able to formulate interventions effectively”, 1-completely disagree, 5-completely agree”.

5) *Therapy Progress*. Assessed with a 10-item Therapy Progress Scale ( $\alpha = 0.85$ ), e.g. “I am satisfied with the overall progress my clients have made during therapy”, 1-completely disagree, 5-completely agree”.

6) *Professional Stress*. Evaluated with the Perceived Stress Scale (Cohen et al., 1983;  $\alpha = 0.79$ ), e.g. “In the last month, how often have you been upset because of something that happened unexpectedly?”, 0-never, 4-very often.

7) *Burnout*. Measured with the Maslach Burnout Inventory (Maslach & Jackson, 1981), which includes three subscales—Emotional Exhaustion, Depersonalization, Personal Accomplishment ( $\alpha = 0.80, 0.75, 0.81$ ), e.g. “I feel emotionally drained by my work”, 0-never, 6-every day.

8) *Job Satisfaction*. Assessed with the modified 10-item Professional Quality of Life Scale (Stamm, 2010;  $\alpha = 0.78$ ), e.g. “I am pleased with how I am able to keep up with my work responsibilities”, 1-never, 5-very often.

For the experimental group, percentage of sessions using Yung Sidekick was also tracked.

### 3.4. Experimental Intervention

Participants in the experimental group used Yung Sidekick, an AI tool designed to automate therapy session note-taking by extracting key session topics, symp-

toms, medications, and goals. The control group continued using their standard documentation practices. The usage of the AI tool by psychotherapists in the experimental group was monitored through the platform's admin system without compromising confidentiality—participants entered a unique code assigned to them at the start of the study.

In addition to self-reports, objective data from the Yung Sidekick admin system were collected. These included timestamps of login and note generation events linked to anonymized user IDs. This metadata enabled verification of session tool usage frequency and time intervals between session end and documentation completion. These logs confirmed that the reported decrease in note-taking time aligned with actual user activity.

Yung Sidekick captures session content through therapist input after the session, either via text summaries or structured prompts. The NLP pipeline processes key clinical terms using transformer-based models (similar to BERT) trained on de-identified therapy transcripts. Extracted items include symptoms, goals, interventions, and medication mentions. All data are encrypted in transit and at rest. The tool adheres to HIPAA compliance standards and includes role-based access, audit logging, and no storage of client identifiers.

### 3.5. Research Design

This study employed a between-subjects randomized controlled trial design with three measurement points: baseline (T0), after two weeks (T1), and after one month (T2).

Factor 1: Use of Yung Sidekick (Experimental vs. Control Group)

Factor 2: Time (T0, T1, T2)

#### *Study Procedures*

1. Baseline Measurement (T0): Participants completed initial assessments and began logging session note-taking and preparation time.

2. Intervention Period (Weeks 1 - 4): The experimental group used Yung Sidekick, while the control group continued standard practices. Both groups tracked documentation and preparation times.

3. Midpoint Measurement (T1): Participants completed the second round of assessments and submitted two-week logs.

4. Final Measurement (T2): Participants completed final assessments and submitted full one-month logs.

#### *Statistical Analysis*

All statistical analyses were conducted using R (Version 4.3.0). Descriptive statistics were computed for demographic variables and outcome measures.

Data was checked for normality using visual (Q-Q plots) and statistical tests (Shapiro-Wilk test). Non-normal data was transformed (e.g., log transformation) to meet the assumptions of parametric tests. Baseline differences between groups were examined using independent samples t-tests for continuous variables and chi-square tests for categorical variables. To handle missing data, multiple imputation

tation (MI) was applied under the assumption that data were missing at random (MAR). Sensitivity analyses were conducted to compare results from different imputation methods.

A Mixed Linear Model (MLM) for repeated measures was used to assess changes over time in key outcome variables. Fixed effects included group assignment, time, and their interaction, while random intercepts were specified for participants to account for individual differences. The model was fitted using the lme4 package, and p-values were obtained via the lmerTest package.

Effect sizes were reported using Cohen's *d* for between-group differences and pseudo  $R^2$  for model fit evaluation. A significance threshold of  $p < 0.05$  was applied across all analyses, and 95% confidence intervals (CI) were reported for key outcome measures.

## 4. Results

**Table 1** presents descriptive statistics for each group across the three measurement points. A comparison of participants in the experimental and control groups at T0 revealed no significant differences in socio-demographic characteristics. Between T0 and T2, the attrition rate was 27.6% in the experimental group and 28.0% in the control group, indicating that engagement levels were comparable across both groups throughout the study period.

Reasons for dropout may include time constraints, technology access issues, and loss of interest (Baumel et al., 2019). We performed both intention-to-treat (including all randomized participants using multiple imputation) and per-protocol analyses (including only those who completed T2). Results from per-protocol analysis showed similar effect directions, suggesting the findings are robust to attrition bias.

**Table 1.** Descriptive statistics.

Scale	Experimental group (n = 21)			Control group (n = 18)		
	T0, M (SD)	T1, M (SD)	T2, M (SD)	T0, M (SD)	T1, M (SD)	T2, M (SD)
Clients per week	20.93 (7.94)	20.50 (7.17)	20.52 (6.07)	21.24 (7.49)	20.72 (7.46)	20.47 (6.62)
Minutes for documenting	20.28 (12.64)	13.17 (7.49)	9.00 (4.59)	20.64 (13.18)	19.89 (11.41)	19.00 (12.07)
Minutes for preparing	14.76 (13.97)	13.29 (12.42)	8.90 (7.47)	14.84 (14.62)	12.56 (10.09)	12.53 (10.68)
Adherence to treatment plan	3.55 (0.69)	3.79 (0.78)	4.10 (0.70)	3.52 (0.71)	3.50 (1.15)	3.40 (1.12)
Therapeutic self-efficacy	4.43 (0.31)	4.35 (0.33)	4.42 (0.33)	4.39 (0.32)	4.37 (0.34)	4.27 (0.38)
Therapy progress	4.02 (0.33)	4.06 (0.36)	4.20 (0.32)	3.98 (0.32)	3.88 (0.42)	3.91 (0.43)
Professional stress	20.76 (3.01)	19.92 (4.75)	18.76 (4.92)	21.48 (3.58)	20.56 (3.35)	18.40 (3.64)
Emotional exhaustion	19.48 (8.81)	20.08 (10.02)	19.38 (10.94)	20.60 (9.69)	19.89 (9.39)	19.13 (8.75)
Depersonalization	4.03 (3.35)	4.42 (3.91)	4.43 (4.42)	4.48 (3.48)	5.78 (3.19)	6.07 (4.10)
Personal accomplishment	38.41(5.77)	39.71 (4.02)	40.67 (4.04)	38.76 (5.87)	36.61 (5.18)	38.07 (6.28)
Job satisfaction	4.10 (0.5)	4.18 (0.41)	4.24 (0.61)	3.98 (0.49)	4.08 (0.48)	4.03 (0.54)

**Table 2.** Results of t-test and mixed linear model analysis.

	t test T0-T2		Mixed Linear Model			pseudo-R <sup>2</sup>
	Experimental group, d	Control group, d	Time effect, $\beta$ (p)	Group effect, $\beta$ (p)	Interaction effect, $\beta$ (p)	
Clients per week	<b>0.043</b>	<b>0.010</b>	-0.034 (0.985)	0.540 (0.752)	-0.715 (0.789)	0.007
Minutes for documenting	<b>1.022***</b>	0.022	<b>-7.653** (0.008)</b>	-0.157 (0.957)	<b>9.811* (.032)</b>	<b>0.164</b>
Minutes for preparing	<b>0.525**</b>	<b>0.050</b>	-5.765 (0.056)	0.018 (0.995)	3.015 (0.512)	0.014
Adherence to treatment plan	<b>0.772**</b>	<b>0.068</b>	<b>0.576* (0.010)</b>	-0.040 (0.856)	<b>-0.679* (0.046)</b>	<b>0.031</b>
Therapeutic self-efficacy	<b>0.006</b>	<b>0.121</b>	-0.090 (.212)	-0.102 (0.167)	-0.052 (0.649)	0.005
Therapy progress	<b>0.479**</b>	<b>0.288</b>	0.170 (0.050)	-0.051 (0.550)	<b>-0.598* (0.048)</b>	<b>0.026</b>
Professional stress	<b>0.326</b>	0.467	-1.384 (0.163)	0.963 (0.318)	-1.660 (0.272)	0.017
Emotional exhaustion	<b>0.018</b>	0.182	1.022 (0.662)	20.831 (0.214)	-4.140 (0.247)	0.010
Depersonalization	<b>0.104</b>	0.251	0.611 (0.530)	0.674 (0.476)	0.464 (0.755)	0.007
Personal accomplishment	<b>0.437*</b>	0.044	1.876 (0.097)	-0.424 (0.700)	-1.874 (0.278)	0.014
Job satisfaction	<b>0.245</b>	0.226	0.139 (0.249)	-0.181 (0.122)	-0.023 (0.902)	0.019

Notes: \*\*\*-p < 0.001, \*\*-p < 0.01, \*-p < 0.05. For Cohen's d, the significance of the Student's t test for dependent samples is given.

The effects of the Yung Sidekick program were assessed using two types of analyses (see **Table 2**). First, paired Student's t-tests were conducted to evaluate within-group changes from T0 to T2, alongside calculations of Cohen's d with 95% confidence intervals (CI) to assess effect sizes. Second, a Mixed Linear Model (MLM) for repeated measures was used to test for time-by-group interaction effects and estimate regression coefficients ( $\beta$ ) with associated CIs and pseudo-R<sup>2</sup> values.

A comparison of T0 - T2 measurements in the experimental group demonstrated a large reduction in session documentation time, from an average of 20.28 minutes to 9.00 minutes (d = 1.022, 95% CI: 0.51 to 1.52), and a moderate reduction in preparation time, from 14.76 minutes to 8.90 minutes (d = 0.525, 95% CI: 0.12 to 0.91). Furthermore, the experimental group showed moderate improvements in adherence to treatment plans (d = 0.779, 95% CI: 0.31 to 1.21), therapy progress (d = 0.479, 95% CI: 0.08 to 0.88), and perceived personal accomplishment (d = 0.437, 95% CI: 0.01 to 0.82). No meaningful changes were observed in the control group across these outcomes.

These results suggest that the intervention yielded moderate-to-large practical effects on key aspects of therapists' workflow and perceived treatment quality. Cohen's d effect size estimates allow us to assess the clinical relevance of these outcomes, providing more informative insight than p-values alone—particularly in the context of a pilot trial with a modest sample size. Cohen's d measures effect size, where d = 0.2 is considered a small effect, d = 0.5 a medium effect, and d = 0.8 or higher a large effect.

The Mixed Linear Model (MLM) for repeated measures further supported these findings. Compared to the control group, the experimental group showed a moderate reduction in documentation time ( $\beta = 9.811$ , 95% CI: 0.91 to 18.71, pseudo- $R^2 = 0.164$ ), a modest increase in adherence to treatment plans ( $\beta = -0.679$ , 95% CI:  $-1.33$  to  $-0.01$ , pseudo- $R^2 = 0.031$ ), and an improvement in therapy progress ( $\beta = -0.598$ , 95% CI:  $-1.19$  to  $-0.01$ , pseudo- $R^2 = 0.026$ ). Other interaction effects were not statistically or practically significant.

Notably, pseudo- $R^2$  values between 0.02 and 0.10 are generally interpreted as small effects, while those between 0.10 and 0.30 indicate moderate effects. In this study, documentation time showed a moderate intervention effect, whereas improvements in adherence and therapy progress reflected small but meaningful effects. These results reinforce the practical utility of AI-assisted tools, especially in improving documentation efficiency.

## 5. Discussion

The findings of this study indicate that the use of Yung Sidekick significantly reduced session documentation time and preparation time among psychotherapists. These reductions, with large and moderate effect sizes respectively, suggest that AI-powered tools can meaningfully improve workflow efficiency in clinical practice. These findings align with previous research demonstrating the benefits of automation in healthcare documentation, where AI-assisted note-taking has been shown to reduce administrative burden and increase time spent in direct patient care (Ghassemi et al., 2020; Lee, Britto, & Diwan, 2024).

Beyond improvements in efficiency, the study revealed significant increases in adherence to treatment plans and perceived therapy progress in the experimental group. However, since these outcomes were measured using self-reported scales, a potential placebo effect should be considered. It is possible that therapists who used an innovative AI tool perceived themselves as more effective simply due to their engagement with a novel technology, rather than actual improvements in therapeutic outcomes. Similar effects have been documented in previous studies on digital interventions, where professionals report increased confidence and effectiveness when using modern technological solutions (Philippe et al., 2022).

Alternatively, these improvements may reflect real changes in therapist behavior. The structured format provided by Yung Sidekick for session notes and treatment plans may have contributed to a more systematic approach to therapy, enhancing therapist adherence to protocols and increasing their overall dedication to client progress. Research has shown that structured documentation tools can promote greater consistency in treatment planning and follow-through (Jensen-Doss et al., 2018). This highlights the potential role of AI in supporting therapists beyond just reducing administrative workload.

Interestingly, no significant changes were observed in other psychological measures such as professional stress, burnout, or job satisfaction. This suggests that while Yung Sidekick may streamline documentation tasks, it does not necessarily alleviate broader professional challenges associated with clinical work. Fu-

ture studies should explore whether prolonged use of AI-assisted tools can contribute to long-term reductions in professional stress or burnout, as seen in some digital interventions designed for healthcare professionals (Jacob, Sanchez-Vazquez, & Ivory, 2020).

Overall, the results of this study provide preliminary evidence that AI-powered documentation tools can enhance therapist efficiency and adherence to treatment plans. However, future research should incorporate objective measures of therapy effectiveness, such as client outcomes or independent assessments of treatment adherence. Additionally, while the observed effects suggest promising benefits of AI-assisted note-taking, further longitudinal studies are necessary to evaluate whether these improvements are sustained over time.

## 6. Limitations

Despite the promising findings, this study has several limitations. First, the reliance on self-reported data introduces potential biases, such as social desirability effects or placebo-like responses from therapists who may perceive themselves as more effective simply due to using an advanced tool. Future studies should incorporate objective measures of adherence to treatment plans and client outcomes. Second, while the sample size was adequate for detecting moderate effects, it remains relatively small for drawing generalizable conclusions. Larger-scale studies are needed to confirm these findings and explore whether AI-assisted note-taking benefits extend to different therapy settings and populations.

Third, the study duration was limited to one month, which does not allow for an assessment of long-term effects. It is unclear whether the observed improvements in efficiency and adherence are sustained over extended periods or if the novelty effect of using AI diminishes over time. Longitudinal studies would provide more insight into the long-term impact of AI tools in therapy practice. Fourth, the study did not control potential differences in the ways therapists interacted with the AI tool. Variability in usage patterns could influence effectiveness, and future research should examine how different engagement levels with AI-assisted documentation impact therapist behavior and client progress.

Finally, generalizability to therapists in other countries remains uncertain, as the study was conducted exclusively with U.S.-based psychotherapists. Cultural and systemic differences in mental health care practices may influence the effectiveness and adoption of AI-assisted tools. Future research should explore the applicability of AI-based documentation solutions across diverse healthcare systems and professional settings. Overall, while this study provides encouraging evidence for the potential of AI-powered note-taking tools in therapy, further research is required to validate these results, mitigate biases, and explore their broader applications in clinical practice.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective User Engagement with Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. *Journal of Medical Internet Research*, *21*, e14567. <https://doi.org/10.2196/14567>
- Budd, J. (2023). Burnout Related to Electronic Health Record Use in Primary Care. *Journal of Primary Care & Community Health*, *14*, 1-7. <https://doi.org/10.1177/21501319231166921>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, *24*, 385-396. <https://doi.org/10.2307/2136404>
- D'Alfonso, S. (2020). AI in Mental Health. *Current Opinion in Psychology*, *36*, 112-117. <https://doi.org/10.1016/j.copsyc.2020.04.005>
- Dawoodbhoy, F. M., Delaney, J., Cecula, P., Yu, J., Peacock, I., Tan, J. et al. (2021). AI in Patient Flow: Applications of Artificial Intelligence to Improve Patient Flow in NHS Acute Mental Health Inpatient Units. *Heliyon*, *7*, e06993. <https://doi.org/10.1016/j.heliyon.2021.e06993>
- Dymek, C., Kim, B., Melton, G. B., Payne, T. H., Singh, H., & Hsiao, C. (2021). Building the Evidence-Base to Reduce Electronic Health Record-Related Clinician Burden. *Journal of the American Medical Informatics Association*, *28*, 1057-1061. <https://doi.org/10.1093/jamia/ocaa238>
- Ebbers, T., Kool, R. B., Smeele, L. E., Dirven, R., den Besten, C. A., Karssemakers, L. H. E. et al. (2022). The Impact of Structured and Standardized Documentation on Documentation Quality; a Multicenter, Retrospective Study. *Journal of Medical Systems*, *46*, Article No. 46. <https://doi.org/10.1007/s10916-022-01837-9>
- Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., Chen, I. Y., & Ranganath, R. (2020). A Review of Challenges and Opportunities in Machine Learning for Health. *AMIA Joint Summits on Translational Science Proceedings, 2020*, 191-200.
- Gori, A., Topino, E., Brugnera, A., & Compare, A. (2022). Assessment of Professional Self-efficacy in Psychological Interventions and Psychotherapy Sessions: Development of the Therapist Self-Efficacy Scale (T-SES) and Its Application for Etherapy. *Journal of Clinical Psychology*, *78*, 2122-2144. <https://doi.org/10.1002/jclp.23391>
- Jacob, C., Sanchez-Vazquez, A., & Ivory, C. (2020). Social, Organizational, and Technological Factors Impacting Clinicians' Adoption of Mobile Health Tools: Systematic Literature Review. *JMIR mHealth and uHealth*, *8*, e15935. <https://doi.org/10.2196/15935>
- Jensen-Doss, A., Haimes, E. M. B., Smith, A. M., Lyon, A. R., Lewis, C. C., Stanick, C. F. et al. (2018). Monitoring Treatment Progress and Providing Feedback Is Viewed Favorably but Rarely Used in Practice. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*, 48-61. <https://doi.org/10.1007/s10488-016-0763-0>
- Lee, C., Britto, S., & Diwan, K. (2024). Evaluating the Impact of Artificial Intelligence (AI) on Clinical Documentation Efficiency and Accuracy across Clinical Settings: A Scoping Review. *Cureus*, *16*, e73994. <https://doi.org/10.7759/cureus.73994>
- Maslach, C., & Jackson, S. E. (1981). The Measurement of Experienced Burnout. *Journal of Organizational Behavior*, *2*, 99-113. <https://doi.org/10.1002/job.4030020205>
- Philippe, T. J., Sikder, N., Jackson, A., Koblanski, M. E., Liow, E., Pilarinos, A. et al. (2022). Digital Health Interventions for Delivery of Mental Health Care: Systematic and Comprehensive Meta-Review. *JMIR Mental Health*, *9*, e35159. <https://doi.org/10.2196/35159>
- Rollwage, M., Habicht, J., Juechems, K., Carrington, B., Viswanathan, S., Stylianou, M. et

al. (2023). Using Conversational AI to Facilitate Mental Health Assessments and Improve Clinical Efficiency within Psychotherapy Services: Real-World Observational Study. *JMIR AI*, 2, e44358. <https://doi.org/10.2196/44358>

Stamm, B. H. (2010). *The Concise ProQOL Manual* (2nd ed.). ProQOL.org.

Zhang, M., Scandiffio, J., Younus, S., Jeyakumar, T., Karsan, I., Charow, R. et al. (2023). The Adoption of AI in Mental Health Care-Perspectives from Mental Health Professionals: Qualitative Descriptive Study. *JMIR Formative Research*, 7, e47847. <https://doi.org/10.2196/47847>